

# Evaluación de Motores de Búsqueda en Internet

Mario Val Perrone – Regina Motz  
Universidad ORT, Montevideo, Uruguay  
e-mail: [mval@iconsulting.com](mailto:mval@iconsulting.com), [motz@athenea.ort.edu.uy](mailto:motz@athenea.ort.edu.uy)

## Resumen

El presente artículo incluye documentación de Motores de Búsqueda obtenida de distintas fuentes y detalla las principales características de ellos, así como también cuáles son los mejores métodos a emplear para la inclusión de sitios web en las bases de datos de estos, de la forma más eficiente.

El principal fundamento y aplicación del presente artículo, será el explicar el funcionamiento de los motores de búsqueda y el servir de guía para el desarrollo de sitios web con el fin que atraigan eficientemente a los motores de búsqueda y Spiders.

## 1 Introducción

La publicación de información en la web es cada día mas voluminosa. El objetivo de quienes publican su información en Internet, ya sea por motivos comerciales o culturales, es que ella llegue al mayor numero de *navegantes* posibles (entendiéndose por navegante, a un usuario de Internet). Sin embargo, el descubrimiento de su sitio por parte del usuario dependerá fundamentalmente de la indexación que sobre el realicen los motores de búsqueda utilizados.

El principal objetivo de este trabajo es evaluar el desempeño de distintos motores de búsqueda de sitios Web con el fin de brindar directivas de diseño para lograr una indexación eficiente, o sea, ser catalogados por la información que consideramos relevante y aparecer en un ranking de alto porcentaje por búsqueda de esa información.

Tomar decisiones fundamentadas a la hora de evaluar los distintos motores de búsqueda y de diseñar un sitio Web para lograr una indexación eficiente es importante considerando que de a cuerdo a datos extraídos de **SearhEngineWach.com**:

- el 75% de los navegantes utilizan motores de búsqueda al ingresar a Internet.
- el 46% de los navegantes, acceden a sitios nuevos a través de motores de búsqueda.
- el 57% de los navegantes realizan al menos 1 búsqueda al día.
- el 42% de los navegantes que realizan compras online, acceden a los sitios de venta a través de motores de búsqueda.
- 1 de cada 28 páginas visualizadas en Internet es el resultado de una búsqueda. (3.5% del total)
- El trabajador promedio que utiliza Internet, pasa 73 minutos al mes en los motores de búsqueda.

El objetivo del artículo es brindar una explicación de cómo funcionan los motores de búsqueda más populares de Internet, brindando datos estadísticos contemporáneos a este documento para justificar, documentar y lograr que el lector tenga un panorama completo de cuáles son las técnicas recomendables a

---

seguir para que un sitio web figure en los motores de búsqueda en los primeros lugares.

El resto del trabajo se organiza de la siguiente manera. En la sección 2 se presentan los motores de búsqueda mas populares y su forma de organizar los sitios. En la sección 3 se presenta una recopilación de cuadros estadísticos de uso de los motores de búsqueda. En la sección 4 discutimos aspectos a tener en cuenta para el diseño de sitios web fácilmente indexables. En la sección 5 mostramos datos experimentales de evaluar la indexación de sitios web con distintos motores de búsqueda. Finalmente en la sección 6 se dan algunas conclusiones.

## **2. Búsqueda de Información**

Generalmente, cuando se inicia una búsqueda de información utilizando un motor de búsqueda, se escribe una o más palabras claves que describen la información que se está buscando. Hay cuatro tipos de opciones para incluir en este tipo de búsquedas:

- **Palabras requeridas y prohibidas.** Para hacer que una palabra esté obligatoriamente dentro de la búsqueda, antecédala de un signo de suma (+). Para hacer que una palabra no se encuentre en la búsqueda, antecédala de un signo de resta (-). Tenga en cuenta que no debe haber espacios entre las palabras y los signos.
- **Restricciones a secciones de documentos.** Para buscar las palabras deseadas sólo en los títulos de los sitios en Internet, utilice "t:" antes de la palabra buscada. Incluir "u:" antes de la palabra buscada limita la búsqueda de la palabra especificada dentro de los enlaces en el web.
- **Búsqueda de Frases.** Para encontrar una frase, utilice las comillas para enmarcar la frase a buscar. No deben existir espacios entre las comillas de apertura y la primera palabra, ni entre la última palabra de la frase y las comillas de cierre.
- **Búsqueda con comodín.** Incluyendo un asterisco (\*) inmediatamente después de la última letra de la palabra buscada, encontrará todas las fuentes que contengan las letras incluidas en la palabra, además de cualquier carácter adicional.
- **Combinaciones.** Se pueden utilizar combinaciones de estas cuatro opciones, solo que debe respetar el mismo orden utilizado para describir cada opción, esto es +, -, t:, u:, " ", y por último \*. Algunos Motores permiten restringir las fechas de los documentos que aparecen en el resultado.
- **Consulta de páginas indexadas de un Host.** Algunos motores de búsqueda, como Altavista, permiten realizar consultas del tipo: **host:undominio.com**, lo cual traerá como resultado todas las páginas que se encuentran indexadas por ejemplo de: undominio.com. Esto permite corroborar cuantas y cuales son las páginas que se encuentran indexadas en dicho motor.

El siguiente paso es esperar a que el motor de búsqueda presente una lista de sitios que coincidan con la petición realizada. Es notorio que los resultados que ofrecen los distintos motores de búsqueda son diferentes. Esto es debido a que funcionan de manera diferente.

---

---

## 2.1 Diferentes Motores de Búsqueda de Información

Hoy en día existen decenas de cientos de motores de búsqueda en Internet, algunos populares y otros no tanto, pero los más populares concentran el mayor volumen de audiencia y de información.

Motores de Búsqueda más populares :

<b>Alta Vista</b>	<a href="http://www.altavista.com/">http://www.altavista.com/</a>
<b>Excite</b>	<a href="http://www.excite.com/">http://www.excite.com/</a>
<b>Hot Bot</b>	<a href="http://www.hotbot.com/">http://www.hotbot.com/</a>
<b>Info Seek</b>	<a href="http://www.infoseek.com/">http://www.infoseek.com/</a>
<b>Lycos</b>	<a href="http://www.lycos.com/">http://www.lycos.com/</a>
<b>Web Crawler</b>	<a href="http://www.webcrawler.com/">http://www.webcrawler.com/</a>
<b>Yahoo!</b>	<a href="http://www.yahoo.com/">http://www.yahoo.com/</a>

Los motores de búsqueda se distinguen por sus métodos para obtener la información:

- Los **sitios de catálogo** mantienen sus propios índices (catálogos) de sitios Web o páginas individuales. Las búsquedas se limitan a los sitios y páginas que están en el catálogo, en lugar de buscarse en todo Internet. (Ejemplos: AltaVista y Lycos.)
- Los **sitios de revisión** elaboran repasos y clasificaciones de sitios Web, y presentan resúmenes temáticos de los sitios que coinciden con su criterio de búsqueda. (Ejemplos: Excite y Magellán.)
- Los **sitios conceptuales** organizan los sitios por categorías y presentan una lista de páginas en categorías relacionadas con su cadena de búsqueda. (Ejemplos: InfoSeek y Yahoo!)

### Motores de Búsqueda vs. Directorios

Si bien comúnmente llamamos "Motores de Búsqueda" tanto a los Motores de Búsqueda como a los Directorios, No son lo mismo.

Los motores de Búsqueda, como Altavista, crean sus índices en forma automática, sin intervención humana, los mismos recorren un sitio y lo indexan y luego los navegantes consultan sobre esos datos recopilados.

Al modificar las páginas, los motores pueden detectarlo y sobrescribir los datos con la nueva información recopilada, ya que los mismos retornan a los sitios en busca de los cambios con frecuencias dadas.

Los directorios como YAHOO! Dependen directamente de humanos para el almacenamiento de información. Quién desee registrar un nuevo sitio, envía a

---

---

través de un formulario web una descripción del sitio, un editor lo revisa y es almacenada. Luego las búsquedas se realizarán sobre dicha descripción. Los cambios que se pueden llegar a hacer para mejorar los rankings en un motor de búsqueda, no son útiles para los Directorios, ya que requieren una revisión manual.

## **2.2 Formas de Indexación y almacenamiento de sitios**

Los motores de búsqueda están compuestos por 3 partes o elementos : **Spider o Crawler, Índice y Software** del Motor de Búsqueda.

El **spider** visita una página, la lee, y continua por los hipervínculos que contenga la página, dentro del sitio. Vuelve a recorrer el sitio según una frecuencia dada a la búsqueda de cambios. (Frecuentemente una vez cada 2 o 4 semanas). Todo lo que el spider encuentra es almacenado en el índice.

El **índice**, algunas veces llamado catálogo, es donde se almacenan las copias de las páginas leídas por el spider. Es posible que un sitio sea recorrido por un spider y no se agregue al índice hasta un momento posterior.

El **Software** del Motor de Búsqueda es el programa encargado de recorrer las millones de páginas almacenadas en el índice, que coinciden con una búsqueda dada, y realizar el ranking de los resultados de acuerdo al grado de relevancia que considere tiene la información.

El Software de cada Motor de Búsqueda se puede comparar con una receta de comida, donde los ingredientes son los mismos, pero cada cocinero, puede agregar sus ingredientes secretos. De esta forma cada motor de búsqueda, si bien debe obtener resultados relevantes a una búsqueda, la forma de determinar cual página o documento indexado es más relevante, varía de un motor a otro. Por ello la búsqueda de una misma cadena de texto en un motor de búsqueda nos da resultados distintos en otros, además de que la información indexada por cada motor es distinta.

Para determinar la relevancia de la información se basan en el principio de Ubicación y Frecuencia.

Lo primero que suelen hacer es buscar que la palabra se encuentre al principio de la página (Título y/o Encabezado de la página o en los primeros 2 párrafos), ya que asumen que lo que se está deseando encontrar, debe figurar desde el principio de una página para que tenga una relevancia considerable. De la misma forma la cantidad de aciertos dentro de la misma página determina cuán relevante puede llegar a ser.

Esto último ha ido evolucionando con el pasar de los años y el aumento de la información, y se han aplicado soluciones a los algoritmos de búsqueda para evitar problemas de spam (repetición de palabras en páginas para obtener mejores rankings en las búsquedas). Hoy en día una gran cantidad de motores de búsqueda son capaces de identificar *spams* y de hacerlo eliminan el sitio del índice.

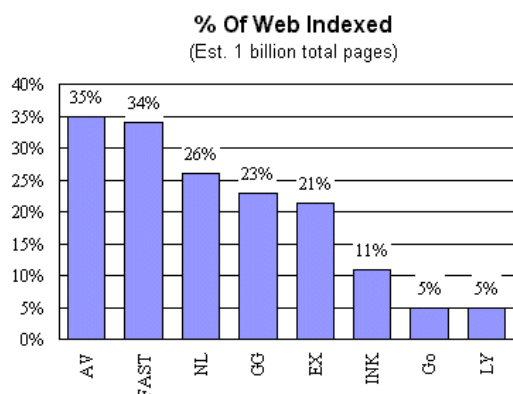
Algunos motores de búsqueda (como Excite) consideran para determinar la relevancia de una página, la cantidad de hipervínculos que existen de otras páginas a ella, ya que consideran que si hay muchas páginas que "apuntan" a esta, se debe tratar de una página importante para dicha búsqueda.

NOTA : En el Anexo I se incluye una tabla con los distintos Spiders utilizados por cada motor de búsqueda.

---

### 3. Algunas Estadísticas

#### Cantidad de Información Indexada vs. Cantidad de Información Existente



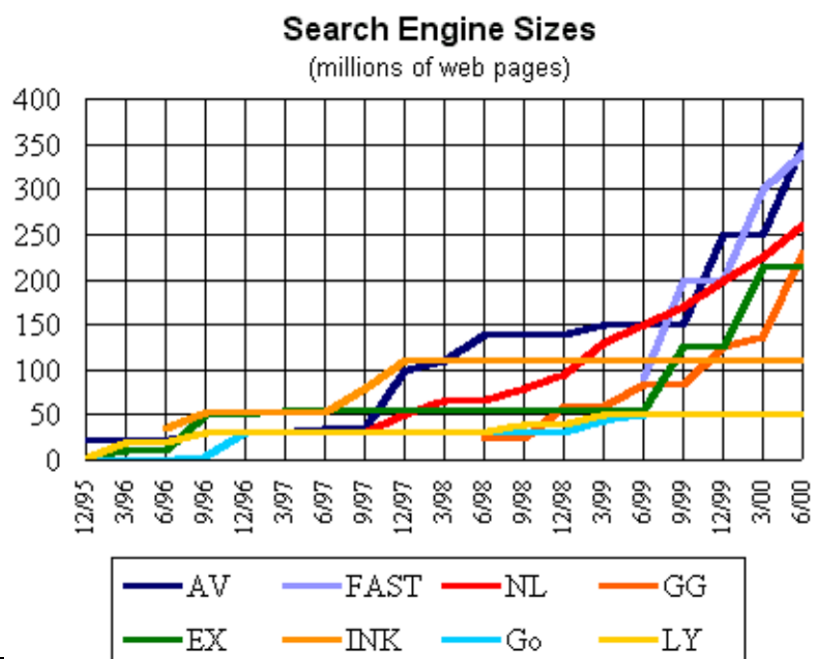
En Febrero del 2000, un estudio en conjunto llevado a cabo por Inktomi y el Departamento de Investigaciones de NEC reveló que existen cerca de 1.000 millones de páginas web indexables en Internet.

La gráfica ilustra como se encuentra “repartida” la Web en cada motor de búsqueda. Podemos apreciar que Altavista es el motor de búsqueda que posee la mayor cantidad de páginas indexadas, las principales razones de esto son :

- Posicionamiento temprano en el mercado, marcando un factor diferenciador al ser de los primeros motores de búsqueda de indexación por contenido.-
- Facilidad para la registración de Páginas.
- Posibilidad de Traducción Online de las páginas.

#### Cantidad de Páginas Indexadas en función del Tiempo

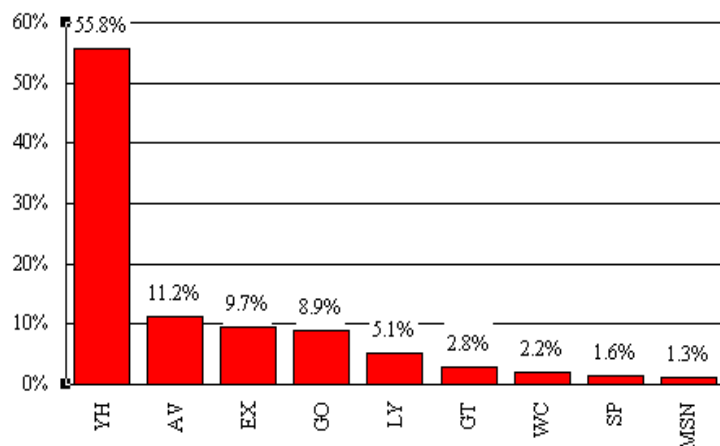
El volumen de información aumenta a diario, de la misma forma que los usuarios que la consultan. La siguiente gráfica ilustra la cantidad de páginas (en millones), almacenadas en los principales motores de búsqueda desde 1996 a la fecha. Se puede apreciar en algunos casos un aumento exponencial año a año en el volumen de páginas indexadas, mientras que en otros como Lycos el volumen se mantiene más conservador.



---

### **Cantidad de páginas vistas a partir de búsquedas realizadas**

El siguiente gráfico ilustra la cantidad de páginas vistas a partir de una búsqueda realizada en un motor de búsqueda en un día dado. Los datos ilustran que sobre la muestra seleccionada, Yahoo fue el motor de búsqueda más utilizado, luego Altavista y Excite en tercer lugar. A pesar de que Yahoo entra dentro de la categoría de sitios conceptuales, y la inclusión de sitios se realiza manualmente demostró ser el más utilizado, por lo cual permite concluir que estos tres motores engloban más del 75% de las páginas buscadas por un navegante.



YH=Yahoo, AV=AltaVista, EX=Excite, GO=Go (Infoseek), LY=Lycos, GT=GoTo, WC=WebCrawler, SP=Snap, MSN=MSN (Información suministrada por StatMarket.)

Los datos fueron recopilados sobre 100.00 sitios web, el día 19 de Diciembre de 1999.

### **4. Algunas pautas de Diseño de sitios Web**

En esta sección analizamos algunos aspectos a tener en cuenta al realizar el diseño del sitio Web para que este sea fácilmente indexable.

#### **Meta Etiquetas**

Las Meta etiquetas o Meta Datos, se encuentran ocultas al navegante, y permiten al diseñador incorporar una breve descripción de lo que el sitio trata y las palabras clave con que el mismo se encuentra relacionado. Algunos Spiders considerarán el contenido de estas etiquetas para aumentar la relevancia de la página. Ejemplos de meta-etiquetas son las siguientes:

**<META name="description" content="una descripción de 200 palabras">**

**<META name="keyword" content="lista de palabras clave separadas por comas">**

Las Meta etiquetas también sirven para indicar a los motores de búsqueda que una página no debe ser indexada por un Spider. Ejemplos de esto son los siguientes:

**<META NAME="robots" CONTENT="noindex">**

Esto indica al Spider que no indexe la página que lo contiene.

**<META NAME="robots" CONTENT="nofollow">**

Esto indica al Spider que indexe la página que lo contiene pero que no continúe indexando las páginas a las cuales está vinculada.

---

---

También se pueden incluir cualquier de las 2 siguientes líneas para que no se indexe la página que la contiene ni a las que estén vinculadas.

**<META NAME="robots" CONTENT="noindex, nofollow">**  
**<META NAME="robots" CONTENT="none">**

Muchos *Webmasters* creen que las meta etiquetas son el secreto para figurar en los primeros lugares en los motores de búsqueda, pues bien, esto no es así. Hotbot e Infoseek consideran mínimamente la existencia de estos a la hora de determinar la relevancia, pero motores como Lycos no los leen en absoluto.

Por lo cual los mismos, son útiles pero no son el ingrediente secreto que va a solucionar los problemas de los *webmasters*, a la hora de figurar en las primeras posiciones.

### **Limitantes existentes**

Existen limitantes a nivel de las Spiders que recorren los sitios alimentándose de lo que encuentran. Muchos Spiders no soportan frames (páginas con cuadros), la única que lo soporta 100% es, el Wolf Spider, de Lycos e Infoseek.

Es importante por ello completar la etiqueta NOFRAME de la página que posea frames, para que al ser recorrida se procese el texto contenido dentro de la misma. Otros tantos Spiders no soportan hipervínculos embebidos en una imagen, únicamente Altavista, Infoseek y Web Crawler soportan mapas de imágenes.

Las páginas generadas dinámicamente, utilizando CGI, no pueden ser indexadas.

## **5. Datos Experimentales**

En esta sección se presentan datos experimentales de prueba de indexación de tres sitios diferentes usando tres motores de búsqueda distintos. El análisis se realiza presentando datos estadísticos referentes a :

- Volumen de Información almacenada.
- Volumen de Audiencia (Visitantes).
- Cantidad de Búsquedas realizadas.
- Soporte de Búsquedas Avanzadas.
- Forma de ejecución de Indexación de Sitios.

Los sitios Web seleccionados para el test fueron los siguientes:

- Promociones y Servicios S.A. [www.psrealestate.com](http://www.psrealestate.com)
- Profesionales SRL [www.profesionalesrsl.com](http://www.profesionalesrsl.com)
- Discount Bank L.A. [www.discbank.com.uy](http://www.discbank.com.uy)

*Promociones y Servicios* es una empresa dedicada a Negocios Inmobiliarios y Administración de Inmuebles.

*Profesionales SRL* es una consultora y financia inversiones principalmente en el sector agropecuario.

*Discount Bank L.A.* es un banco subsidiario del Israel Discount Bank of New York.

Los tres sitios fueron registrados en los motores de Búsqueda :

- Altavista.
  - Yahoo!
  - Lycos.
-

---

Fecha de la registraci3n : **5 de Junio del 2000.**

Fecha de las consultas : **10 de Julio del 2000.**

### **ALTAVISTA**

A continuaci3n de detallar3n las b3squedas realizadas para cada uno de los dominios seleccionados en Altavista.

host: psrealestate.com - Trajo como resultado 21 p3ginas.  
host: profesionalesrsl.com - Trajo como resultado 2 p3ginas.  
host: discbank.com.uy - Trajo como resultado 41 p3ginas.

Real Estate Brokers Uruguay - Primer resultado.  
Real Estate Brokers Montevideo - Primer resultado.  
Real Estate Brokers Punta del Este - Primer resultado.  
+Alquiler +Propiedades +Uruguay 5ta. posici3n  
"Promociones y Servicios" - Primer resultado.  
+Promociones +Servicios - Primer resultado.  
"Punta del Este" - 12va. posici3n

" +Financiamiento +Inversiones" - Primer lugar : [www.profesionalesrsl.com.uy](http://www.profesionalesrsl.com.uy)  
"Profesionales" - No figura  
"Profesionales Srl" - No figura  
"Profesionales SRL" - Primer Lugar  
+Profesionales +Uruguay - No figura

Discount Bank - 2do. lugar  
"Discount" (S3lo Espa3ol) - Primer lugar p3gina principal  
[www.discbank.com.uy](http://www.discbank.com.uy)  
"+Bank +Latin +America" - Primer lugar : [www.discbank.com.uy](http://www.discbank.com.uy) de :  
39.134.

### **YAHOO!**

A continuaci3n de detallar3n las b3squedas realizadas para cada uno de los dominios seleccionados en Yahoo!

www.psrealestate.com - Trajo como resultado 7 p3ginas.  
www.profesionalesrsl.com - Trajo como resultado 2 p3ginas.  
www.discbank.com.uy - Trajo como resultado 5 p3ginas.

+Alquiler +Propiedades +Uruguay 8 resultados sin figurar.  
"Promociones y Servicios" - Primer resultado.  
+Promociones +Servicios - Primer resultado.  
Real Estate Brokers Uruguay- No figura.  
Real Estate Brokers Montevideo - No figura.  
Real Estate Brokers Punta del Este -No figura.

" +Financiamiento +Inversiones" - Primer lugar : [www.profesionalesrsl.com.uy](http://www.profesionalesrsl.com.uy)  
"Profesionales" - No figura  
"Profesionales srl" - Primer resultado  
"Profesionales SRL" - Primer resultado

---



---

+Profesionales +Uruguay	-	No figura (Existen 320 resultados)
Discount Bank	-	No figura.
"Discount" (Sólo Español)	-	No figura.
" +Bank +Latin +America"	-	No figura.

## **LYCOS**

A continuación se detallarán las búsquedas realizadas para cada uno de los dominios seleccionados en Lycos.

host:psrealestate.com	-	Trajo como resultado 11 páginas.
host:profesionalesrl.com	-	Trajo como resultado 5 páginas.
host:discbank.com.uy	-	Trajo como resultado 12 páginas.

+Alquiler +Propiedades +Uruguay	-	No figura.
"Promociones y Servicios"	-	3er- 6to. 7mo. Lugar
+Promociones +Servicios	-	3er. 5to. Y 7mo. Lugar.
Real Estate Brokers Uruguay-		Segundo Lugar.
Real Estate Brokers Montevideo -		Segundo Lugar.
Real Estate Brokers Punta del Este	-	Segundo Lugar.

" +Financiamiento +Inversiones"	-	Primer Lugar.
+Profesionales +srl	-	Los primeros 15 resultados pertenecen al sitio.
"Profesionales SRL"	-	Los primeros 15 resultados pertenecen al sitio.
+Profesionales +Uruguay	-	No figura (Existen 28451 resultados)
" +Bank +Latin +America"	-	Primeros 3 resultados pertenecen al sitio-
"Bancos Uruguay"	-	No figura en 9314 resultados.
Discount Bank	-	4to. Lugar.

## **6. Conclusiones**

A partir de los datos recogidos se puede afirmar claramente las diferencias existentes en los distintos componentes de los motores de búsqueda y las distintas políticas de indexación.

Por ello debemos tener en cuenta a la hora de diseñar un Sitio y armar las páginas que lo compongan ciertos puntos que pueden ser claves a la hora de figurar en las bases de datos de los motores de búsqueda indexados por las palabras que se desea.

Altavista, a diferencia de los restantes, es sensitivo a mayúsculas y minúsculas. Los algoritmos para la búsqueda y determinación de relevancia varían en las prioridades dadas a los siguientes aspectos:

- título de la página.
- metadatos definidos.
- combinación de los dos anteriores.
- repetición de las palabras clave.

Algunos motores no emplean los metadatos definidos para indexar.

---

---

Altavista indexa las páginas a partir de los primeros 1024 caracteres de la página, tomando de esa forma el título, encabezados y metaetiquetas existentes en cada página.

Los algoritmos empleados por Lycos definen la relevancia a partir de la cantidad de aciertos de una búsqueda en el título de la página y los metadatos definidos en ella.

Yahoo!, al agrupar los sitios en categorías en forma manual, es más lento en lo que refiere a la indexación y en uno de los casos analizados no figura, luego de haber solicitado la indexación, un mes atrás.

Los motores de búsqueda, continuarán mejorando sus componentes, con el avance de la tecnología, permitiendo a los navegantes encontrar lo que estos buscan en forma cada vez más eficiente.

Los medios de almacenamiento de información, al igual que los algoritmos empleados para la búsqueda e indexación de datos continuarán evolucionando, surgirán nuevas tecnologías que hoy en día se encuentran en fase de prototipo o prueba, las cuales permitirán ir resolviendo los "open problems" existentes.

De todas formas en lo que respecta al presente, existen algunos puntos que se pueden considerar para lograr una indexación correcta por parte de los motores búsqueda a la hora de publicar un sitio web.

Estos puntos son :

- Hacer buen uso de las etiquetas de las páginas.
- Armar eficazmente el encabezado de cada página.
- Utilizar Meta-etiquetas.
- No hacer páginas con excesiva cantidad de texto.
- No utilizar Frames o Cuadros.
- Ser paciente.

## **Referencias**

En los siguientes sitios Web podrá encontrar información complementaria al presente informe.

Search Engine Watch	-	<a href="http://www.searchenginewatch.com">www.searchenginewatch.com</a>
FAST Search	-	<a href="http://www.ussc.alltheweb.com">http://www.ussc.alltheweb.com</a>
Advanced Searching	-	<a href="http://www.onlineinc.com/onlinemag/MayOL/zorn5.html">http://www.onlineinc.com/onlinemag/MayOL/zorn5.html</a>
Kansas City Public Library	-	<a href="http://www.kclibrary.org/">http://www.kclibrary.org/</a>
Search Engine Spiders	-	<a href="http://www.siteware.ch/webresources/useragents/spiders/">http://www.siteware.ch/webresources/useragents/spiders/</a>
Search Engine Spiders II	-	<a href="http://www.u.arizona.edu/~iia/spider.html">http://www.u.arizona.edu/~iia/spider.html</a>
Search Engine Features	-	<a href="http://webreference.com/content/search/features.html">http://webreference.com/content/search/features.html</a>
WWW9 Papers	-	<a href="http://searchenginewatch.com/sereport/00/06-tech.html">http://searchenginewatch.com/sereport/00/06-tech.html</a>

# ANEXO I

## Gráfica de Comparación entre Motores de Búsqueda

	AltaVista	Excite	HotBot	Infoseek	Lycos	Northern Light
<b>Tamaño</b>	250 Millones de páginas	250 Millones de páginas y objetos multimedia	110 Millones de Sitios	75 Millones de páginas	50 Millones de páginas	200 Millones de Sitios
<b>Full-text</b>	SI	SI	SI	SI	NO	SI
<b>Palabra por defecto</b>	OR	OR	AND	or	and	and
<b>Boolean connectors</b>	AND, AND NOT, NEAR	AND, AND NOT	OR, NOT	and, not	or, not, adj., near, before, far	or, not
<b>Búsqueda de Frase</b>	Usar Comillas	Usar Comillas	Usar Comillas	Usar Comillas	Usar Comillas	Usar Comillas
<b>Truncado</b>	No, usar *	No	No, ? y *	No	No	Automático
<b>Sensible a Mayúsculas</b>	SI	No	SI	SI	No	No
<b>Inclusión de Palabras</b>	Usar +	Usar +	Usar +	Usar +	Usar +	Usar +
<b>Eliminación de Palabras</b>	Usar -	Usar -	Usar -	Usar -	Usar -	Usar -
<b>Detección de Duplicados</b>	Agupados por un mismo título	SI	Agupados por un mismo título	SI	SI	SI
<b>Especiales</b>	Limitar por fecha, lenguaje o un campo seguido por una coma	Búsqueda por Conceptos	Limitar por fecha, lenguaje, ubicación y profundidad de página	Encontrar búsquedas similares	Buscar por imagines y archivos de sonido	Carpetas personalizadas

Información suministrada por :



Kansas City  
Public Library

**A continuación se detallan los distintos agentes o Spiders que utilizan los motores de búsqueda más populares :**

<b>Motor de Búsqueda</b>	<b>Nombre del Agente</b>	<b>Nombre del Host</b>
<a href="#">AltaVista</a> (normal spider)	Scooter/2.0 G.R.A.B. X2.0 Scooter/1.0 scooter@pa.dec.com	scooter.pa-x.dec.com scooter*.av.pa-x.dec.com ejs.: scooter3.av.pa-x.dec.com
<a href="#">AltaVista</a> (instant spider)	Scooter/1.0	add-url.altavista.digital.com ww2.altavista.digital.com
<a href="#">Excite</a> (mega spider)	ArchitextSpider	crawl*.atext.com ej. : crawl2.atext.com
<a href="#">Excite</a> (fresh spider)	ArchitextSpider	crimpshrine.atext.com
<a href="#">Google</a> (Experimental search engine)	BackRub/2.1 backrub@google.stanford.edu http://google.stanford.edu/	*.stanford.edu ej. : hake.stanford.edu
<a href="#">Inktomi</a> (powers <a href="#">HotBot</a> , others)	Slurp/2.0 (slurp@inktomi.com; http://www.inktomi.com/slurp.html)	*.inktomi.com ejs : j2001.inktomi.com or j10.inktomi.com
<a href="#">Infoseek</a> (normal spider)	InfoSeek Sidewinder/0.9	*.infoseek.com ejs: wilbur-bbn.infoseek.com o IP: 204.162.98.90
<a href="#">Infoseek</a> (instant spider)	Mozilla/3.01 (Win95; I)	Idem anterior
<a href="#">Lycos</a> (regular spider)	Lycos_Spider_(T-Rex)	lycosidae.lycos.com o *.pgh.lycos.com ejs : spider3.srv.pgh.lycos.com
<a href="#">Lycos</a> (Add URL spider)	Lycos_Spider_(T-Rex)	*.sjc.lycos.com ej: sjc-fe4-1.sjc.lycos.com
<a href="#">WebCrawler</a>	Utiliza Excite	Servido por Excite

---

**Cantidad de Búsquedas realizadas en un mes.**

Motor	Cantida de Búsquedas	Fecha de Informado
AltaVista	55 millones	05/2000
Inktomi	47 millones	04/2000
Google	12 millones	04/2000
Ask Jeeves	4 millones	03/2000
Voila	1.5 millones	01/2000

FUENTE : **SearchEngineWatch.** - <http://www.searchenginewatch.com/reports/perday.html>

**La siguiente tabla ilustra las preferencias de los navegantes a la hora de consultar diferentes motores de búsquedas**

Servicio	AV	AJ	EX	GO	GT	HB	LS	LY	MSN	NS	SP	YH
AV	x	24%	22%	23%	25%	25%	29%	26%	19%	16%	20%	18%
AJ	4%	x	4%	4%	4%	5%	5%	3%	3%	3%	4%	3%
EX	29%	31%	x	23%	38%	33%	30%	30%	20%	30%	35%	20%
GO	43%	36%	33%	x	35%	42%	36%	41%	32%	30%	40%	30%
GT	15%	14%	17%	11%	x	20%	25%	19%	11%	19%	21%	9%
HB	17%	20%	17%	15%	22%	x	21%	27%	11%	15%	18%	12%
LS	23%	20%	18%	15%	33%	24%	x	22%	13%	21%	20%	13%
LY	36%	26%	31%	30%	44%	55%	39%	x	26%	29%	37%	23%
MSN	52%	42%	42%	47%	53%	44%	45%	52%	x	39%	46%	40%
NS	29%	31%	43%	30%	59%	41%	51%	40%	26%	x	47%	27%
SP	14%	17%	19%	15%	25%	19%	18%	19%	12%	18%	x	11%
YH	68%	65%	58%	63%	61%	67%	64%	65%	57%	56%	62%	

---

---

	Alta Vista	InfoSeek	Lycos
<b># de URLs</b>	16 millones	1 millón	10,75 millones
<b>Documentación</b>	Excelente; incluye los detalles y los ejemplos de la búsqueda para las interrogaciones simples y avanzadas.	Muy largo; algo difícil de localizar.	Justo, tan completo como podía ser.
<b>Detección Duplicada</b>	No	Sí	No
<b>El Buscar Del Campo</b>	Sí	No	No
<b>Indexación de direcciones</b>	Con texto completo	Con texto completo	URLs, también opción de otras partes de paginaciones y texto.
<b>Conjuntos Múltiples De la Búsqueda</b>	No	No	No
<b>Boleano Jerarquizada</b>	Sí	No	No
<b>El Buscar De la Proximidad</b>	Sí, asegura ambos términos están dentro de diez palabras (o de frases si los términos se incluyen en cotizaciones).	Sí	Sí, elija el tipo del emparejamiento (flojo, feria, los etc.) de menú pull-down.
<b>Graduación De la Importancia</b>	Sí, puede especificar qué términos al peso primero en la visualización de la extracción.	Sí	Sí
<b>Truncamiento</b>	Sí, pueden utilizar un asterisco en vástagos de la palabra más de largo de tres cartas.	No	Sí, automático.

---